

On Random Operator-Valued Matrices: Operator-Valued Semicircular Mixtures and Central Limit Theorem

Mario Diaz*

October 15, 2014

Abstract

Motivated by a random matrix theory model from wireless communications, we define random operator-valued matrices as the elements of $L^\infty(\Omega, \mathcal{F}, \mathbb{P}) \otimes M_d(\mathcal{A})$ where $(\Omega, \mathcal{F}, \mathbb{P})$ is a classical probability space and (\mathcal{A}, φ) is a non-commutative probability space. A central limit theorem for the mean $M_d(\mathbb{C})$ -valued moments of these random operator-valued matrices is derived. Also a numerical algorithm to compute the mean $M_d(\mathbb{C})$ -valued Cauchy transform of operator-valued semicircular mixtures is analyzed.

Keywords: random operator-valued matrices, central limit theorem, operator-valued free probability, operator-valued semicircular mixture, multiantenna system model.

1 Introduction

In this section we motivate the study of what we call random operator-valued matrices. In particular, we generalize a random matrix model used in wireless communications. This will also provide us a natural link between random matrix ensembles and random operator-valued matrices. In the following section we summarize the notation and setting of this paper.

Recent developments in operator-valued free probability theory [1, 2, 3] have made possible to analyze a variety of wireless communication systems [4, 5, 6]. In the developing area of massive multiantenna systems [7] the dimension of the random matrix modelling the system is in the order hundreds or thousands. These large sizes suggest that the behavior of the spectrum of these matrices is very close to their asymptotic models, e.g. free deterministic equivalents [5] and operator-valued equivalents [6].

Roughly speaking, random operator-valued matrices (models) is a special class of random variables with values over the matrices with coefficients in some non-commutative algebra. This contrasts with the classical models studied before, e.g. [5, 6], where the asymptotic models are non-random operator-valued matrices over some non-commutative algebra. From an applied point of view, this extra randomness may reflect

*Department of Mathematics and Statistics, Queen's University, Kingston, ON, Canada, 13madt@queensu.ca

the statistical variations of the channel in a scale of time bigger than a period of use. Therefore this kind of model may be relevant to study properties of channels that depend on statistics that change over large periods of time.

Random operator-valued matrices are also a natural object from a mathematical point of view. It is known that if operator-valued matrices, i.e. elements in $M_d(\mathcal{A})$ where (\mathcal{A}, φ) is a non-commutative probability space, have entries free over \mathcal{A} then they are free over $M_d(\mathbb{C})$. Thus, if we have two fixed families $\{a_{i,j}\}_{i,j=1}^d$ and $\{b_{i,j}\}_{i,j=1}^d$ free over \mathcal{A} and we construct operator-valued matrices $\mathbf{A} = f(a_{i,j}; i, j)$ and $\mathbf{B} = g(b_{i,j}; i, j)$ for some suitable functions $f, g : \mathcal{A}^{d^2} \rightarrow M_d(\mathcal{A})$, then \mathbf{A} and \mathbf{B} will be free over $M_d(\mathbb{C})$. If instead of a pair of functions we have a pair of families of functions $\{f_\omega\}_{\omega \in \Omega}$ and $\{g_\omega\}_{\omega \in \Omega}$ for some probability space $(\Omega, \mathcal{F}, \mathbb{P})$, then $\mathbf{A} = \mathbf{A}(\omega) = f_\omega(a_{i,j}; i, j)$ and $\mathbf{B} = \mathbf{B}(\omega) = g_\omega(b_{i,j}; i, j)$ are random variables with values over operator-valued matrices, thus the name random operator-valued matrices. Of course some measurability conditions should be satisfied, but in the context of this paper this requirement will be clearly satisfied. If the families $\{f_\omega\}_{\omega \in \Omega}$ and $\{g_\omega\}_{\omega \in \Omega}$ are suitable then the freeness is preserved almost everywhere, which makes possible to work in the realm of operator-value free probability theory almost everywhere. This is similar to classical probability where we work in the realm of real analysis almost everywhere. In this sense, here we are dealing with a probabilistic version of operator-value free probability theory. We will not work at this too general level of abstraction, in what follows we construct a particular class of objects for which the pointwise operator-valued behavior is more structured.

Returning to the wireless communication context, the random matrix of interest H_N have the following form

$$H_N = \begin{pmatrix} A_{1,1}X_N^{(1,1)} & \cdots & A_{1,d}X_N^{(1,d)} \\ \vdots & \ddots & \vdots \\ A_{d,1}X_N^{(d,1)} & \cdots & A_{d,d}X_N^{(d,d)} \end{pmatrix}$$

where $A = (A_{i,j})_{i,j=1}^d$ is a selfadjoint $d \times d$ random matrix and $\{X_N^{(i,j)} \mid 1 \leq i, j \leq d\}$ is a family of independent $N \times N$ standard complex Gaussian matrices such that $X_N^{(j,i)} = (X_N^{(i,j)})^*$ for all $1 \leq i, j \leq d$. The asymptotic analysis will be done with respect to N while d remain fixed (see [6] for further details on this kind of models). The matrix A is assumed to be independent of $X_N^{(i,j)}$ for all $1 \leq i, j \leq d$ and such that all its entries belong to $L^\infty(\Omega, \mathcal{F}, \mathbb{P})$. In fact, the extra randomness added to this model with respect to the one analyzed in [6] comes from this matrix A .

Let F^{H_N} be empirical eigenvalue distribution of H_N . Some of the quantities of interest in the wireless communication context are given by

$$I_N := \mathbb{E} \left(\int_{-\infty}^{\infty} f(x) dF^{H_N}(x) \right)$$

for some non-negative, continuous and bounded function f . By a standard argument we have then

$$I_N = \int_{-\infty}^{\infty} f(x) dF_N(x)$$

where F_N is the mean eigenvalue distribution of H_N , i.e. $F_N(x) = \mathbb{E}(F^{H_N}(x))$ for all $x \in \mathbb{R}$.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be the underlying probability space. Suppose that the previous hypothesis on H_N are satisfied¹ for all $N \in \mathbb{N}$, then for almost every $\omega \in \Omega$ we have that [8]

$$H_N(\omega) \xrightarrow{\text{dist}} \begin{pmatrix} A_{1,1}(\omega)\mathbf{X}_{1,1} & \cdots & A_{1,d}(\omega)\mathbf{X}_{1,d} \\ \vdots & \ddots & \vdots \\ A_{d,1}(\omega)\mathbf{X}_{d,1} & \cdots & A_{d,d}(\omega)\mathbf{X}_{d,d} \end{pmatrix} =: \mathbf{H}(\omega) \quad (1)$$

where $\{\mathbf{X}_{i,j} \mid 1 \leq i, j \leq d\}$ is a free circular family in a non-commutative probability space (\mathcal{A}, φ) with $\mathbf{X}_{i,j}^* = \mathbf{X}_{j,i}$ for all $1 \leq i, j \leq d$ and thus $\mathbf{X} = (\mathbf{X}_{i,j})_{i,j=1}^d \in M_d(\mathcal{A})$ is an operator-valued matrix. Observe that $\mathbf{H}(\omega) = A(\omega) \circ \mathbf{X}$ where \circ denotes the Hadamard or entrywise product. By convergence in distribution we mean that

$$\lim_{N \rightarrow \infty} (\mathbf{I} \otimes \text{tr}_N)(H_N(\omega)^m) \stackrel{\text{a.s.}}{=} E(\mathbf{H}(\omega)^m)$$

for all $m \in \mathbb{N}$ where tr_N is the normalized trace in $M_N(\mathbb{C})$ and $E := \mathbf{I} \otimes \varphi$. Up to this point we described the behavior of $\mathbf{H}(\omega)$ for a fixed $\omega \in \Omega$. We can propose several abstract spaces in which the expression $\mathbf{H} := A \circ \mathbf{X}$ have sense, in the next paragraph we construct such a space and describe the relations that the expectation w.r.t. \mathbb{P} , \mathbb{E} , and E should have.

To find the intended relations first consider a $N \times N$ block of the matrix H_N^m for $m \in \mathbb{N}$, say the i_1, i_m -block for some $i_1, i_m \in \{1, \dots, d\}$. Such a block is indeed a random matrix, so the natural linear functional to study is $\mathbb{E} \circ \text{tr}_N$. The block under study is the sum of random matrices of the form

$$A_{i_1, i_2} \cdots A_{i_{m-1}, i_m} X_N^{(i_1, i_2)} \cdots X_N^{(i_{m-1}, i_m)}$$

with $i_2, \dots, i_{m-1} \in \{1, \dots, d\}$. The independence between A and the matrices $\{X_N^{(i,j)} \mid 1 \leq i, j \leq d\}$ implies that

$$\begin{aligned} \lim_{N \rightarrow \infty} \mathbb{E} \left(\text{tr}_N \left(A_{i_1, i_2} \cdots A_{i_{m-1}, i_m} X_N^{(i_1, i_2)} \cdots X_N^{(i_{m-1}, i_m)} \right) \right) &= \mathbb{E} (A_{i_1, i_2} \cdots A_{i_{m-1}, i_m}) \varphi (\mathbf{X}_{i_1, i_2} \cdots \mathbf{X}_{i_{m-1}, i_m}) \\ &= (\mathbb{E} \otimes \varphi)(A_{i_1, i_2} \cdots A_{i_{m-1}, i_m} \otimes \mathbf{X}_{i_1, i_2} \cdots \mathbf{X}_{i_{m-1}, i_m}). \end{aligned}$$

This suggests that the entries of \mathbf{H} may belong to $(L^{\infty-}(\Omega, \mathcal{F}, \mathbb{P}), \mathbb{E}) \otimes (\mathcal{A}, \varphi)$. Therefore, \mathbf{H} can be thought as an element in the space $(L^{\infty-}(\Omega, \mathcal{F}, \mathbb{P}), \mathbb{E}) \otimes (M_d(\mathcal{A}), E)$. From the algebraic construction of

¹One way to construct such a family is: a) construct a probability space where the random matrix A exists, b) construct another probability space where the family of $\{X_N^{(i,j)} \mid N \in \mathbb{N}, 1 \leq i, j \leq N\}$ exists and c) in the product of these spaces take the inclusions associated to aforementioned random variables.

the previous tensor product, $\mathbb{E} \otimes 1_{\mathcal{A}}$ and $1 \otimes \varphi$ commute when applied to elements in $L^{\infty-}(\Omega, \mathcal{F}, \mathbb{P}) \otimes \mathcal{A}$, $(1 \otimes E)(A) = A$ for all random matrix A and $(\mathbb{E} \otimes \text{I})(\mathbf{X}) = \mathbf{X}$ for all \mathbf{X} (non-random) operator-valued matrix and thus $1 \otimes E$ and $\mathbb{E} \otimes \text{I}$ commute. If \mathcal{B} is a sub σ -algebra of \mathcal{F} , from the definition of conditional expectation we have that $(\mathbb{E} \otimes 1_{\mathcal{A}})(\mathbb{E}(A \mid \mathcal{B}) \otimes x) = (\mathbb{E} \otimes 1_{\mathcal{A}})(A \otimes x)$ for $A \in L^{\infty-}(\Omega, \mathcal{F}, \mathbb{P})$ and $x \in \mathcal{A}$, and in particular taking conditional expectation with respect to \mathcal{B} (in the first coordinate) and applying $1 \otimes \varphi$ commute in $L^{\infty-}(\Omega, \mathcal{F}, \mathbb{P}) \otimes \mathcal{A}$.

By the independence between A and the matrices $\{X_N^{(i,j)} \mid 1 \leq i, j \leq d\}$, the convergence in equation (1) also holds in mean and thus

$$\lim_{N \rightarrow \infty} \mathbb{E}(\text{tr}_{dN}(H_N^m)) = \mathbb{E}(\text{tr}_d(E(\mathbf{H}^m)))$$

for all $m \in \mathbb{N}$, or equivalently

$$\lim_{N \rightarrow \infty} \int_{\mathbb{R}} t^m dF_N(t) = \int_{\mathbb{R}} t^m dF(t)$$

for all $m \in \mathbb{N}$, where F is the mean analytical distribution of \mathbf{H} (see Definition 10 and Proposition 1). If F is determined by its moments, the convergence of the moments implies that $F_N \Rightarrow F$ and in particular $I_N \rightarrow I := \int_{\mathbb{R}} f(t) dF(t)$. Therefore we can take I as an approximation for I_N and focus on the mean analytical distribution F of \mathbf{H} .

The variance of the entries of X can be absorbed by A , so without loss of generality we assume that $\varphi(\mathbf{X}_{i,j} \mathbf{X}_{k,l}) = \delta_{i,l} \delta_{j,k}$. Likewise, if $A_{i,k} = |A_{i,k}| \exp(i \arg(A_{i,k}))$ then the equality in distribution as complex random matrices $\exp(i \arg(A_{i,k})) X_N^{(i,k)} \stackrel{d}{=} X_N^{(i,k)}$ and the independence between A and the family $\{X_N^{(i,j)} \mid 1 \leq i, j \leq d\}$ show that the argument of $A_{i,k}$ can be absorbed by $\mathbf{X}_{i,k}$, so without loss of generality we assume that the entries of A are positive random variables.

It is important to remark that the random operator-valued matrix of the form $A \circ \mathbf{X}$ as described above can be thought as the mixture by A of the operator-valued semicircular element (over $M_d(\mathbb{C})$) \mathbf{X} . Due to the important role of the classical Gaussian mixture in several applications, this analogy provides a general motivation for the study of elements of the aforementioned form.

2 Notation and Setting

Throughout this paper d will be a fixed positive integer. $M_d(\mathbb{C})$ denotes the $d \times d$ matrices over the complex numbers and $D_d(\mathbb{C})$ the subset of diagonal complex matrices. The notation $A_{i,k}$ or $(A)_{i,k}$ denotes the i, k -entry of the matrix A . We set $M_d(\mathbb{C})^{\pm} = \{B \in M_d(\mathbb{C}) \mid 0 < \pm \Im(B)\}$ where $\Im(B) := \frac{B - B^*}{2i}$, and in particular $\mathbb{C}^{\pm} = M_1(\mathbb{C})^{\pm}$. Also, \circ denotes the pointwise or Hadamard product of matrices.

In what follows $(\Omega, \mathcal{F}, \mathbb{P})$ will be a probability space with expectation \mathbb{E} . We will denote by (\mathcal{A}, φ) a non-commutative probability space where $1_{\mathcal{A}}$ is the unit in \mathcal{A} . The algebra $(M_d(\mathcal{A}), E)$ where $E := \text{I} \otimes \varphi : M_d(\mathcal{A}) \equiv M_d(\mathbb{C}) \otimes \mathcal{A} \rightarrow M_d(\mathbb{C})$ is called the operator-valued matrices. The algebra of random operator-valued matrices is then defined to be $(L^{\infty-}(\Omega, \mathcal{F}, \mathbb{P}), \mathbb{E}) \otimes (M_d(\mathcal{A}), E)$. Abusing of the notation, we also use

E to denote the map $1 \otimes (\mathbf{I} \otimes \varphi) : L^{\infty-}(\Omega, \mathcal{F}, \mathbb{P}) \otimes M_d(\mathcal{A}) \rightarrow M_d(\mathbb{C})$. In the same spirit, we use \mathbb{E} to denote both expectation in $(\Omega, \mathcal{F}, \mathbb{P})$ and the map $\mathbb{E} \otimes \mathbf{I} : L^{\infty-}(\Omega, \mathcal{F}, \mathbb{P}) \otimes M_d(\mathcal{A}) \rightarrow M_d(\mathcal{A})$. We will use lower case letters to denote scalars, upper case letters to denote both matrices and random matrices, and bold upper case letters to denote both operator-valued matrices and random operator-valued matrices.

3 Definitions

Recall the following definitions from [9].

Definition 1. Let $\mathbf{X}^{(1)}$ and $\mathbf{X}^{(2)}$ be operator-valued elements in $M_d(\mathcal{A})$. We say that $\mathbf{X}^{(1)}$ and $\mathbf{X}^{(2)}$ are free over $M_d(\mathbb{C})$ if

$$E\left(p_1(\mathbf{X}^{(i_1)})p_2(\mathbf{X}^{(i_2)})\cdots p_k(\mathbf{X}^{(i_k)})\right) = 0$$

whenever $k \in \mathbb{N}$, $i_j \neq i_{j+1}$ for all $1 \leq j < k$ and $\{p_i \mid 1 \leq i \leq k\}$ are non commutative polynomials over $M_d(\mathbb{C})$ such that $E(p_j(\mathbf{X}^{(i_j)})) = 0$ for all $1 \leq j \leq k$.

Definition 2. We say that $\mathbf{X} \in M_d(\mathcal{A})$ is an operator-valued semicircular element over $M_d(\mathbb{C})$ if

$$E(\mathbf{X}^m) = \sum_{\pi \in \text{NC}_2(m)} \kappa_{\pi}(\mathbf{X})$$

where $\text{NC}_2(m)$ are the non-crossing pairings of the set $\{1, \dots, m\}$ and $\kappa_{\pi}(\mathbf{X})$ is defined recursively using the nested structure of π and the function $\eta : M_d(\mathbb{C}) \rightarrow M_d(\mathbb{C})$ given by $\eta(B) = E(\mathbf{X}B\mathbf{X})$ for $B \in M_d(\mathbb{C})$.

Example 1. In the notation of the previous definition, if $\pi = \{(1, 6), (2, 3), (4, 5), (7, 8)\}$ then

$$\kappa_{\pi}(\mathbf{X}) = \eta(\eta(\mathbf{I})\eta(\mathbf{I}))\eta(\mathbf{I}).$$

Definition 3. We define an operator-valued semicircular mixture, semicircular mixture for short, as a random operator-valued matrix \mathbf{H} such that $\mathbf{H} = A \circ \mathbf{X}$ where A is a $d \times d$ selfadjoint random matrix with non-negative entries in $L^{\infty-}(\Omega, \mathcal{F}, \mathbb{P})$ and \mathbf{X} is a selfadjoint operator-valued matrix such that $\{\mathbf{X}_{i,j} \mid 1 \leq i, j \leq d\}$ is a free circular family up to symmetry conditions².

Recall that any random operator-valued matrix \mathbf{Z} , i.e. an element in $L^{\infty-}(\Omega, \mathcal{F}, \mathbb{P}) \otimes M_d(\mathcal{A})$, can be written as $\sum_{i=1}^n a_i \otimes \mathbf{Z}_i$ for some $n \in \mathbb{N}$ and $a_1, \dots, a_n \in L^{\infty-}(\Omega, \mathcal{F}, \mathbb{P})$ and $\mathbf{Z}_1, \dots, \mathbf{Z}_n \in M_d(\mathcal{A})$.

Definition 4. Let $\mathbf{Z}^{(1)}$ and $\mathbf{Z}^{(2)}$ be two elements in $L^{\infty-}(\Omega, \mathcal{F}, \mathbb{P}) \otimes M_d(\mathcal{A})$. We say that $\mathbf{Z}^{(1)}$ and $\mathbf{Z}^{(2)}$ are independent w.r.t. \mathbb{P} if there exist $n_1, n_2 \in \mathbb{N}$, $a_1^{(1)}, \dots, a_{n_1}^{(1)}, a_1^{(2)}, \dots, a_{n_2}^{(2)} \in L^{\infty-}(\Omega, \mathcal{F}, \mathbb{P})$ and $\mathbf{Z}_1^{(1)}, \dots, \mathbf{Z}_{n_1}^{(1)}, \mathbf{Z}_1^{(2)}, \dots, \mathbf{Z}_{n_2}^{(2)} \in M_d(\mathcal{A})$ such that $\mathbf{Z}^{(1)} = \sum_{i=1}^{n_1} a_i^{(1)} \otimes \mathbf{Z}_i^{(1)}$, $\mathbf{Z}^{(2)} = \sum_{i=1}^{n_2} a_i^{(2)} \otimes \mathbf{Z}_i^{(2)}$ and $\{a_1^{(1)}, \dots, a_{n_1}^{(1)}\}$ and $\{a_1^{(2)}, \dots, a_{n_2}^{(2)}\}$ are independent families of random variables. Similarly, we say that $\mathbf{Z}^{(1)}$

²In particular, the condition $\mathbf{X}_{i,i} = \mathbf{X}_{i,i}^*$ implies that the diagonal elements are semicircular non-commutative random variables.

and $\mathbf{Z}^{(2)}$ are free over $M_d(\mathbb{C})$ if the families $\{\mathbf{Z}_1^{(1)}, \dots, \mathbf{Z}_{n_1}^{(1)}\}$ and $\{\mathbf{Z}_1^{(2)}, \dots, \mathbf{Z}_{n_2}^{(2)}\}$ are free over $M_d(\mathbb{C})$.

The linearity of E implies that if $\mathbf{Z}^{(1)}$ and $\mathbf{Z}^{(2)}$ are free over $M_d(\mathbb{C})$ (in the sense of the previous definition), for any non-commutative polynomial in two variables p the expression $E(p(\mathbf{Z}^{(1)}, \mathbf{Z}^{(2)}))$ equals to the same expression as if $\mathbf{Z}^{(1)}$ and $\mathbf{Z}^{(2)}$ were two (non-random) operator-valued elements free over $M_d(\mathbb{C})$, e.g. $E(\mathbf{Z}^{(1)}\mathbf{Z}^{(2)}\mathbf{Z}^{(1)}) = E(\mathbf{Z}^{(1)}E(\mathbf{Z}^{(2)})\mathbf{Z}^{(1)})$.

Definition 5. Let \mathbf{Z} be a random operator-valued matrix. We say that \mathbf{Z} is centered if there exists $n \in \mathbb{N}$, $a_1, \dots, a_n \in L^{\infty-}(\Omega, \mathcal{F}, \mathbb{P})$ and $\mathbf{Z}_1, \dots, \mathbf{Z}_n \in M_d(\mathcal{A})$ such that $\mathbf{Z} = \sum_{i=1}^n a_i \otimes \mathbf{Z}_i$ and $E(\mathbf{Z}_i) = 0$ for all $1 \leq i \leq n$.

Observe that centered is used to refer just a property of the operator-valued part. The reason for doing this is that indeed we will not need any centeredness assumption for the random part. This reflects the fact that the operator-valued part *dominates* the random part in terms of the impact to the overall behavior of a random operator-valued matrix.

Recall the following standard definitions.

Definition 6. The scalar-valued Cauchy transform $g_F : \mathbb{C}^+ \rightarrow \mathbb{C}^-$ of a probability distribution F is defined for $z \in \mathbb{C}^+$ by

$$g_F(z) = \int_{\mathbb{R}} \frac{1}{z - t} dF(t).$$

Definition 7. Let $\mathbf{X} \in M_d(\mathcal{A})$ be an operator-valued matrix, we define its $M_d(\mathbb{C})$ -valued Cauchy transform $G_{\mathbf{X}}(B)$ for $B \in M_d(\mathbb{C})$ such that $B - \mathbf{X}$ is invertible by

$$G_{\mathbf{X}}(B) = E((B - \mathbf{X})^{-1}).$$

Following the pointwise or almost sure philosophy discussed early, the next definition is the straightforward generalization of the previous ones to our random operator-valued context.

Definition 8. Let $\mathbf{H} \in L^{\infty-}(\Omega, \mathcal{F}, \mathbb{P}) \otimes M_d(\mathcal{A})$ be a random operator-valued matrix, we define its mean $M_d(\mathbb{C})$ -valued Cauchy transform $\mathcal{G}_{\mathbf{H}}$ by

$$\mathcal{G}_{\mathbf{H}}(B) = \mathbb{E}(G_{\mathbf{H}}(B))$$

for $B \in M_d(\mathbb{C})$ such that $B - \mathbf{H}$ is invertible almost surely.

Observe that $G_{\mathbf{H}}(B)$ is random matrix as it depends on \mathbf{H} . It is also possible to write $\mathcal{G}_{\mathbf{H}}(B)$ as $(\mathbb{E} \otimes E)((B - \mathbf{H})^{-1})$. However, we prefer the form used in the previous definition since, as we will see throughout the paper, the behavior of E *dominates* the one of \mathbb{E} .

Given a selfadjoint $\mathbf{X} \in M_d(\mathcal{A})$, the function $h : \mathbb{C}^+ \rightarrow \mathbb{C}$ defined by $h(z) = \text{tr}_d(G_{\mathbf{X}}(zI))$ satisfies that $h(\mathbb{C}^+) \subset \mathbb{C}^-$ and $\lim_{y \rightarrow \infty} iyh(iy) = 1$. Therefore [9], there exists a unique probability distribution $F^{\mathbf{X}}$ on \mathbb{R}

such that

$$h(z) = \int_{\mathbb{R}} (z - t)^{-1} dF^{\mathbf{X}}(t).$$

This observation gives sense to the following definition and its natural generalization to the random operator-valued context.

Definition 9. Let $\mathbf{X} \in M_d(\mathcal{A})$ be selfadjoint. We define the analytical distribution of \mathbf{X} as the unique probability distribution $F^{\mathbf{X}}$ on \mathbb{R} such that

$$\mathrm{tr}_d(G_{\mathbf{X}}(z\mathbf{I})) = \int_{\mathbb{R}} (z - t)^{-1} dF^{\mathbf{X}}(t)$$

for $z \in \mathbb{C}^+$.

Definition 10. Let $\mathbf{H} \in L^{\infty-}(\Omega, \mathcal{F}, \mathbb{P}) \otimes M_d(\mathcal{A})$ be a selfadjoint random operator-valued matrix. We define the mean analytical distribution of \mathbf{H} as the probability distribution F on \mathbb{R} such that $F(x) := \mathbb{E}(F^{\mathbf{H}}(x))$ where $F^{\mathbf{H}(\omega)}$ is the analytical distribution of $\mathbf{H}(\omega)$.

4 Mean $M_d(\mathbb{C})$ -Valued Cauchy Transform of Semicircular Mixtures

Let $\mathbf{H} = A \circ \mathbf{X}$ be a semicircular mixture and F its mean analytical distribution. Observe that $F^{\mathbf{H}}$ is a random probability distribution on \mathbb{R} as it depends on A . The definition of F (see Definition 10) requires averaging (w.r.t. \mathbb{P}) the analytical distributions $F^{\mathbf{H}}$. The next proposition shows that averaging these distributions and then taking the Cauchy transform is the same as averaging the corresponding Cauchy transforms. Additionally, this proposition proves that this is also true for the moments, i.e. the moments of the mean analytical distribution of \mathbf{H} are equal to the mean scalar-valued moments of \mathbf{H} , and they exist. In particular, this shows that the mean analytical distribution is the right object to study the behavior of \mathbf{H} .

Proposition 1. Let $\mathbf{H} = A \circ \mathbf{X}$ be a semicircular mixture and F its mean analytical distribution. Then the scalar-valued Cauchy transform $g_F : \mathbb{C}^+ \rightarrow \mathbb{C}$ of F is given by

$$g_F(z) = \mathrm{tr}_d(\mathcal{G}_{\mathbf{H}}(z\mathbf{I})).$$

For all $m \in \mathbb{N}$ the m -th moment of F exists and

$$\int_{\mathbb{R}} t^m dF(t) = \mathbb{E}(\mathrm{tr}_d(E(\mathbf{H}^m))).$$

The previous proposition imply that g_F can be obtained from the mean $M_d(\mathbb{C})$ -valued Cauchy transform $\mathcal{G}_{\mathbf{H}}$ of \mathbf{H} . By the Stieltjes inversion theorem, it is enough to compute $\mathcal{G}_{\mathbf{H}}$ to obtain F . Therefore we will focus on $\mathcal{G}_{\mathbf{H}}$ in what follows.

Definition 11. Let $\mathbf{H} = A \circ \mathbf{X}$ be a semicircular mixture. We define the random mapping $\eta : D_d(\mathbb{C}) \rightarrow$

$D_d(\mathbb{C})$ for $D \in D_d(\mathbb{C})$ by

$$\eta(D)_{i,i} = \sum_{j=1}^d A_{i,j}^2 D_{j,j}$$

for all $1 \leq i \leq d$.

Theorem 1. Let $\mathbf{H} = A \circ \mathbf{X}$ be a semicircular mixture. Then

$$\mathcal{G}_{\mathbf{H}}(z\mathbf{I}) = \lim_{n \rightarrow \infty} \mathbb{E}(T_{z\mathbf{I}}^{\circ n}(W))$$

for all $z \in \mathbb{C}^+$ and all $W \in M_d(\mathbb{C})^-$ where $T_{z\mathbf{I}}(D) = (z\mathbf{I} - \eta(D))^{-1}$ for $D \in D_d(\mathbb{C})$.

The previous theorem is a straightforward extension of the operator-valued version in [1]. In particular, this extension shows that limit and expectation commute. This commutativity implies that taking expectation or limit first does not matter when computing $\mathcal{G}_{\mathbf{H}}$. The latter suggests that any reasonable numerical method will approximate $\mathcal{G}_{\mathbf{H}}$ robustly. In particular, the following routine is an example of such a method.

By definition, for $D \in D_d(\mathbb{C})$ we have that $T_{z\mathbf{I}}^{\circ n+1}(D) = T_{z\mathbf{I}}(T_{z\mathbf{I}}^{\circ n}(D))$, that is

$$T_{z\mathbf{I}}^{\circ n+1}(D)_{i,i} = \left(z - \sum_{j=1}^d A_{i,j}^2 T_{z\mathbf{I}}^{\circ n}(-i\mathbf{I})_{j,j} \right)^{-1} \quad (2)$$

for $i \in \{1, \dots, d\}$. So we can approximate $\mathcal{G}_{\mathbf{H}}(z\mathbf{I})$ using Monte-Carlo method as follows

- Put $S = \mathbf{0} \in D_d(\mathbb{C})$ and iterate M times the following subroutine;
 - Pick random values for A ;
 - Define $T_{z\mathbf{I}}^{\circ 0} = -i\mathbf{I}$;
 - Compute $T_{z\mathbf{I}}^{\circ n+1}$ from $T_{z\mathbf{I}}^{\circ n}$ as in equation (2);
 - Stop at $n = N$ and add $T_{z\mathbf{I}}^{\circ N}$ to S ;
- Approximate $\mathcal{G}_{\mathbf{H}}(z\mathbf{I})$ by S/M .

Observe that the above procedure have complexity $O(MNd^2)$. To speed up the algorithm some tailor made modifications were done. See the next section for details.

At a first sight the previous routine may seem quite unsatisfactory as it depends on Monte-Carlo method. However, suppose that we have a general formula for $\mathcal{G}_{\mathbf{H}}$ that depends on the joint distribution of the entries of A . If such a formula comes in the form of an integral with respect to this joint distribution then $\mathcal{G}_{\mathbf{H}}(z)$ would be an integral over a $d(d+1)/2$ -dimensional space. Even for d relatively small, such an integral is likely to be evaluated by a Monte-Carlo-like method. Our point here is that in general the previous routine is as far as we can go numerically speaking.

We also have the following corollaries.

Corollary 1. Let $\mathbf{H} = A \circ \mathbf{X}$ be a semicircular mixture. If for each $\omega \in \Omega$ there exists $K(\omega)$ such that

$K(\omega) = \sum_{j=1}^d A_{i,j}^2$ for all $i \in \{1, \dots, d\}$, then $\mathcal{G}_{\mathbf{H}}(z\mathbf{I})$ is given by

$$\mathcal{G}_{\mathbf{H}}(z\mathbf{I}) = \mathbb{E} \left(\frac{z - \sqrt{z^2 - 4K}}{2K} \right) \mathbf{I}.$$

In particular, if K does not depend on ω then $\mathcal{G}_{\mathbf{H}}(z\mathbf{I})$ is given by $\mathcal{G}_{\mathbf{H}}(z\mathbf{I}) = \frac{z - \sqrt{z^2 - 4K}}{2K} \mathbf{I}$.

The entries of A were supposed to be positive for simplicity, but nothing stop us from using the results found so far to any matrix as long as the squares are replaced by the appropriate square norms.

Corollary 2. Let $\mathbf{H} = A \circ \mathbf{X}$ be a semicircular mixture and F its mean analytical distribution. If A is a unitary selfadjoint random matrix then $F = F^{\mathbf{X}}$.

In the particular case when A is constant the previous corollary is a remarkable, and already known, property of semicircular elements over $M_d(\mathbb{C})$. In this sense, this corollary asserts that the random version of this known fact is also true. This makes reasonable that the structure of the moments of the CLT limit in $L^{\infty-}(\Omega, \mathcal{F}, \mathbb{P}) \otimes M_d(\mathcal{A})$ is connected to those in the operator-valued case.

5 Central Limit Theorem

Suppose that $\mathbf{Z} \in L^{\infty-}(\Omega, \mathcal{F}, \mathbb{P}) \otimes M_d(\mathcal{A})$ can be written as $\sum_{i=1}^n a_i \otimes \mathbf{Z}_i$ for some $n \in \mathbb{N}$, $a_1, \dots, a_n \in L^{\infty-}(\Omega, \mathcal{F}, \mathbb{P})$ and $\mathbf{Z}_1, \dots, \mathbf{Z}_n \in M_d(\mathcal{A})$. We construct $\{\mathbf{Z}^{(j)}\}_j$ independent and free over $M_d(\mathbb{C})$ copies of \mathbf{Z} by a) creating $\{a_1^{(j)}, \dots, a_n^{(j)}\}_j$ independent identically distributed copies of $\{a_1, \dots, a_n\}$, b) creating $\{\mathbf{Z}_1^{(j)}, \dots, \mathbf{Z}_n^{(j)}\}_j$ free over $M_d(\mathbb{C})$ copies of $\{\mathbf{Z}_1, \dots, \mathbf{Z}_n\}$ and c) defining $\mathbf{Z}^{(j)} = \sum_{i=1}^n a_i^{(j)} \otimes \mathbf{Z}_i^{(j)}$.

We now establish a central limit theorem for random operator-valued matrices. The proof of the following central limit theorem is analogous to the operator-valued case, though it shows that independence, freeness over $M_d(\mathbb{C})$ and centeredness assumptions³ imply that the limiting mean $M_d(\mathbb{C})$ -moments have the same structure as the $M_d(\mathbb{C})$ -moments of semicircular element in $M_d(\mathcal{A})$. This is a quite remarkable feature as the non-trivial extension $L^{\infty-}(\Omega, \mathcal{F}, \mathbb{P}) \otimes M_d(\mathcal{A})$ of $M_d(\mathcal{A})$ have the same structure for the mean $M_d(\mathbb{C})$ -moments of the CLT limit as the CLT limit in $M_d(\mathcal{A})$ itself.

Theorem 2. Let $\mathbf{Z} \in L^{\infty-}(\Omega, \mathcal{F}, \mathbb{P}) \otimes M_d(\mathcal{A})$ be a selfadjoint centered random operator-valued matrix. Suppose that $\{\mathbf{Z}^{(n)}\}_{n \geq 1}$ are independent and free over $M_d(\mathbb{C})$ copies of \mathbf{Z} . Then the normalized sum $\mathbf{S}_N = \frac{\mathbf{Z}^{(1)} + \dots + \mathbf{Z}^{(N)}}{\sqrt{N}}$ satisfies

$$\lim_{N \rightarrow \infty} \mathbb{E}(E(\mathbf{S}_N^m)) = \sum_{\pi \in \text{NC}_2(m)} \hat{\kappa}_{\pi}(\mathbf{Z})$$

for all $m \in \mathbb{N}$ where $\hat{\kappa}(\pi)$ is defined as in Definition 2 using $\hat{\eta}(B) = \mathbb{E}(E(\mathbf{Z}B\mathbf{Z}))$ ($B \in M_d(\mathbb{C})$) instead of η .

³In the sense of definitions 4 and 5.

Here is worth to point out the following. Consider $d = 1$ and the non-commutative probability space $(L^{\infty-}(\Omega, \mathcal{F}, \mathbb{P}) \otimes \mathcal{A}, \hat{\varphi} := \mathbb{E} \otimes \varphi)$. Suppose a_1, a_2 are independent symmetric Bernoulli random variables in $L^{\infty-}(\Omega, \mathcal{F}, \mathbb{P})$ and $b_1, b_2 \in \mathcal{A}$ are free (w.r.t. φ) Poisson non-commutative random variables. By definition, $a_1 \otimes b_1$ and $a_2 \otimes b_2$ are independent and free over $M_1(\mathbb{C})$. Observe that

$$\begin{aligned} \hat{\varphi}((a_1 \otimes b_1)(a_2 \otimes b_2)(a_1 \otimes b_1)(a_2 \otimes b_2)) &= \mathbb{E}(a_1^2 a_2^2) \varphi(b_1 b_2 b_1 b_2) \\ &= \varphi(b_1^2) \varphi(b_2)^2 + \varphi(b_1)^2 \varphi(b_2^2) - \varphi(b_1)^2 \varphi(b_2)^2 = 7, \end{aligned}$$

but $\hat{\varphi}((a_1 \otimes b_1)^2) \hat{\varphi}(a_2 \otimes b_2)^2 + \hat{\varphi}(a_1 \otimes b_1)^2 \hat{\varphi}((a_2 \otimes b_2)^2) - \hat{\varphi}(a_1 \otimes b_1)^2 \hat{\varphi}(a_2 \otimes b_2)^2 = 0$ as $\mathbb{E}(a_1) = \mathbb{E}(a_2) = 0$ and so $\hat{\varphi}(a_1 \otimes b_1) = \hat{\varphi}(a_2 \otimes b_2) = 0$. In particular, $a_1 \otimes b_1$ and $a_2 \otimes b_2$ are not free w.r.t. $\hat{\varphi}$. Therefore the previous theorem is a different result from the operator-valued central limit theorem, even when it gives us the same structure for the $M_d(\mathbb{C})$ -moments.

As a particular consequence of the central limit theorem derived we have the following.

Corollary 3. Let $\{A^{(n)}\}_{n \geq 1}$ be i.i.d. $d \times d$ selfadjoint random matrices with common distribution A with entries in $L^{\infty-}(\Omega, \mathcal{F}, \mathbb{P})$ and let $\{\mathbf{X}^{(n)}\}_{n \geq 1}$ be selfadjoint operator-valued centered elements free over $M_d(\mathbb{C})$ with common distribution \mathbf{X} . Then the normalized sum $\mathbf{S}_N = \frac{A^{(1)} \circ \mathbf{X}^{(1)} + \dots + A^{(N)} \circ \mathbf{X}^{(N)}}{\sqrt{N}}$ satisfies

$$\lim_{N \rightarrow \infty} \mathbb{E}(E(\mathbf{S}_N^m)) = \mathbb{E}(E(\mathbf{Y}^m))$$

for all $m \in \mathbb{N}$ where $\mathbf{Y} \in M_d(\mathcal{A})$ is an operator-valued semicircular element with covariance given by

$$\varphi(\mathbf{Y}_{i,k} \mathbf{Y}_{j,l}) = \mathbb{E}(A_{i,k} A_{j,l}) \varphi(\mathbf{X}_{i,k} \mathbf{X}_{j,l}). \quad (3)$$

Observe that in the previous corollary we do not assume any particular distribution for \mathbf{X} .

6 Numerical Example

To illustrate the technique discussed after Theorem 1, we compute the following example. We take the empirical eigenvalue distribution of 1000 matrices of size 300×300 with the following distribution: $d = 3$ and $N = 100$; the operator-valued part \mathbf{X} was approximated by $(W + W^*)/20$ where W is a 300×300 complex Gaussian random matrix with i.i.d. entries with 0 mean and variance 2; the random part A was taken as a 3×3 matrix with i.i.d. standard Rayleigh distributed entries, up to symmetries. To approximate $\mathcal{G}_{\mathbf{H}}(z\mathbf{I})$ we used the Monte-Carlo technique with 1000 iterations. To speed up the iterative method it did not stop at a fixed N but when the difference between two consecutive iterations was smaller than 10^{-3} in $\|\cdot\|_1$. To make the iterative method even faster, given a realization of A , we computed $G_{\mathbf{H}}(z\mathbf{I})$ for all the desired values of z , say $z[1], \dots, z[L]$, using as initial value to iterate $T_{z[n+1]\mathbf{I}}$ the last value $T_{z[n]\mathbf{I}}^{\circ N}$.

We took $\Im(z[n]) = 0.001$. In the author experience, this value is small enough to provide a good

approximation, as showed in the following picture, and at the same time is far enough from the real axis to ensure convergence of the iterative method. Needless to say, the figure shows good agreement between the observed eigenvalues and the estimated density computed from $\mathcal{G}_{\mathbf{H}}$.

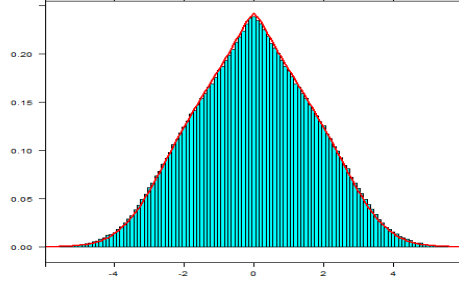


Figure 1: Histogram of the eigenvalues and plot of the estimated theoretical density.

7 Proofs of the Main Results

Proof of Proposition 1. It is well known that

$$\int_{\mathbb{R}} f(t) dF(t) = \mathbb{E} \left(\int_{\mathbb{R}} f(t) dF^{\mathbf{H}}(t) \right)$$

for all positive or F -integrable function f . Let $z \in \mathbb{C}^+$ be fixed. By definition $g_F(z) = \int_{-\infty}^{\infty} \frac{1}{z-t} dF(t)$. Since $|z-t|^{-1} \leq \Im(z)^{-1}$, the function $(z-t)^{-1}$ is bounded and in particular integrable w.r.t. F . Thus

$$g_F(z) = \mathbb{E} \left(\int_{-\infty}^{\infty} \frac{1}{z-t} dF^{\mathbf{H}}(t) \right).$$

By definition of analytical distribution, the integral inside the expectation is equal to $\text{tr}_d(G_{\mathbf{H}}(z\mathbf{I}))$ and therefore

$$\begin{aligned} g_F(z) &= \mathbb{E}(\text{tr}_d(G_{\mathbf{H}}(z\mathbf{I}))) \\ &= \text{tr}_d(\mathbb{E}(G_{\mathbf{H}}(z\mathbf{I}))) \\ &= \text{tr}_d(\mathcal{G}_{\mathbf{H}}(z\mathbf{I})), \end{aligned}$$

where the middle equality follows from the fact that trace and expected value commute.

For $\omega \in \Omega$ fixed the values of $\{A_{i,j} \mid 1 \leq i, j \leq d\}$ are also fixed and finite a.s., since these are the variances of the entries of $\mathbf{H}(\omega)$ we have then that the support of $F^{\mathbf{H}(\omega)}$ is compact and also $\text{tr}_d(E(\mathbf{H}(\omega)^m)) =$

$\int_{\mathbb{R}} t^m dF^{\mathbf{H}(\omega)}(t)$ for all $m \in \mathbb{N}$ [9]. Therefore, for $m \in \mathbb{N}$,

$$\begin{aligned}\mathbb{E}(\text{tr}_d(E(\mathbf{H}^m))) &= \mathbb{E}\left(\int_{\mathbb{R}} t^m dF^{\mathbf{H}}(t)\right) \\ &= \int_{\mathbb{R}} t^m dF(t),\end{aligned}$$

where the last equation, as in the previous paragraph, is true as long as t^m is integrable w.r.t. F . We will prove that t^m is integrable w.r.t. F .

Let $m \in \mathbb{N}$ be fixed. Since $|t|^m$ is positive,

$$\int_{\mathbb{R}} |t|^m dF(t) = \mathbb{E}\left(\int_{\mathbb{R}} |t|^m dF^{\mathbf{H}}(t)\right) \leq \mathbb{E}(\lambda_{\max}(\mathbf{H})^m)$$

where $\lambda_{\max}(\mathbf{H})(\omega)$ is the supremum of the absolute value over the support of $F^{\mathbf{H}(\omega)}$, i.e.

$$\lambda_{\max}(\mathbf{H})(\omega) = \sup_{\lambda \in \text{Supp}(F^{\mathbf{H}(\omega)})} |\lambda|.$$

If we find a constant $C(\omega)$ such that $\int_{\mathbb{R}} t^n dF^{\mathbf{H}(\omega)}(t) \leq C(\omega)^n$ for all $n \in \mathbb{N}$, then we will have that $\lambda_{\max}(\mathbf{H})(\omega) \leq C(\omega)$. Recall that $\int_{\mathbb{R}} t^n dF^{\mathbf{H}(\omega)}(t) = \text{tr}_d(E(\mathbf{H}(\omega)^n))$ for $n \in \mathbb{N}$ and $\omega \in \Omega$ fixed. A straightforward computation shows that for $n \in \mathbb{N}$

$$\begin{aligned}\int_{\mathbb{R}} t^n dF^{\mathbf{H}(\omega)}(t) &= \text{tr}_d(E(\mathbf{H}(\omega)^n)) = \frac{1}{d} \sum_{i_1=1}^d \varphi\left((\mathbf{H}(\omega)^n)_{i_1, i_1}\right) \\ &= \frac{1}{d} \sum_{i_1, i_2, \dots, i_n=1}^d \varphi(\mathbf{H}(\omega)_{i_1, i_2} \cdots \mathbf{H}(\omega)_{i_n, i_1}) \\ &= \frac{1}{d} \sum_{i_1, i_2, \dots, i_n=1}^d (A_{i_1, i_2} \cdots A_{i_n, i_1})(\omega) \varphi(\mathbf{X}_{i_1, i_2} \cdots \mathbf{X}_{i_n, i_1}).\end{aligned}$$

The Wick type formula for a free circular family shows that $\varphi(\mathbf{X}_{i_1, i_2} \cdots \mathbf{X}_{i_n, i_1}) \geq 0$ for all $i_1, \dots, i_n \in \{1, \dots, d\}$. Let $M(\omega) = \max_{i, j=1, \dots, d} A_{i, j}(\omega)$. By assumption $A_{i, j}$ is positive for all $1 \leq i, j \leq d$, so the positivity of the coefficients $\varphi(\mathbf{X}_{i_1, i_2} \cdots \mathbf{X}_{i_n, i_1})$ implies

$$\begin{aligned}\int_{\mathbb{R}} t^n dF^{\mathbf{H}(\omega)}(t) &\leq \frac{1}{d} \sum_{i_1, i_2, \dots, i_n=1}^d M(\omega)^n \varphi(\mathbf{X}_{i_1, i_2} \cdots \mathbf{X}_{i_n, i_1}) \\ &= M(\omega)^n \text{tr}_d(E(\mathbf{X}^n)).\end{aligned}$$

Since the analytical distribution of \mathbf{X} is compact [9], we have that $\lambda_{\max}(\mathbf{X}) := \sup_{\lambda \in \text{Supp}(F^{\mathbf{X}})} |\lambda|$ is finite

and $\text{tr}_d(E(\mathbf{X}^n)) \leq \lambda_{\max}(\mathbf{X})^n$. Thus

$$\int_{\mathbb{R}} t^n dF^{\mathbf{H}(\omega)}(t) \leq M(\omega)^n \lambda_{\max}(\mathbf{X})^n,$$

and in particular $\lambda_{\max}(\mathbf{H}) \leq \lambda_{\max}(\mathbf{X})M$. Therefore

$$\begin{aligned} \int_{\mathbb{R}} |t|^m dF(t) &\leq \mathbb{E}(\lambda_{\max}(\mathbf{H})^m) \\ &\leq \lambda_{\max}(\mathbf{X})^m \mathbb{E}(M^m) < \infty, \end{aligned}$$

where the existence of the m -th moment of the maximum is guaranteed by the fact that $A_{i,j} \in L^{\infty-}(\Omega, \mathcal{F}, \mathbb{P})$ for all $1 \leq i, j \leq d$. \square

Recall the following theorem from [1]. We rephrase it in our terminology, so it constitutes the basis for the pointwise analysis (w.r.t. $\omega \in \Omega$).

Theorem 3. Let $\mathbf{H} = A \circ \mathbf{X}$ be a semicircular mixture. Fix $\omega \in \Omega$. Define the mapping $\eta : M_d(\mathbb{C}) \rightarrow M_d(\mathbb{C})$ given by $\eta(B) = E(\mathbf{H}(\omega)B\mathbf{H}(\omega))$. Then, for $B \in M_d(\mathbb{C})^+$, the $M_d(\mathbb{C})$ -valued Cauchy transform of $\mathbf{H}(\omega)$ is given by

$$G_{\mathbf{H}(\omega)}(B) = \lim_{n \rightarrow \infty} T_B^{\circ n}(W) \quad (4)$$

for any $W \in M_d(\mathbb{C})^-$ where $T_B(W) := (B - \eta(W))^{-1}$. Moreover, $G_{\mathbf{H}(\omega)}(B) \in M_d(\mathbb{C})^-$ and satisfies

$$(B - \eta(G_{\mathbf{H}(\omega)}(B)))G_{\mathbf{H}(\omega)}(B) = I. \quad (5)$$

Before proving Theorem 1 we need to prove the following lemmas. The next lemma shows that the definition of η in Definition 11 actually coincides with the definition of η in the previous theorem for semicircular mixtures.

Lemma 1. Let $\mathbf{H} = A \circ \mathbf{X}$ be a semicircular mixture and let $\eta : M_d(\mathbb{C}) \rightarrow M_d(\mathbb{C})$ be the random mapping defined by $\eta(B) = E(\mathbf{H}B\mathbf{H})$ for $B \in M_d(\mathbb{C})$. If $B \in M_d(\mathbb{C})$ then $\eta(B) \in D_d(\mathbb{C})$ and

$$\eta(B)_{i,i} = \sum_{j=1}^d A_{i,j}^2 B_{j,j}.$$

Proof. Let $B \in M_d(\mathbb{C})$ and $i, k \in \{1, \dots, d\}$. By definition of E we have that $\eta(B)_{i,k} = \varphi((\mathbf{H}B\mathbf{H})_{i,k})$ and

a straightforward computation shows that

$$\begin{aligned}\eta(B)_{i,k} &= \sum_{i_1, i_2=1}^d \varphi(\mathbf{H}_{i, i_1} B_{i_1, i_2} \mathbf{H}_{i_2, k}) \\ &= \sum_{i_1, i_2=1}^d A_{i, i_1} A_{i_2, k} B_{i_1, i_2} \varphi(\mathbf{X}_{i, i_1} \mathbf{X}_{i_2, k}).\end{aligned}$$

From the fact that $\{\mathbf{X}_{i,k} \mid 1 \leq i, k \leq d\}$ is a free circular family with $\mathbf{X}_{i,j} = (\mathbf{X}_{j,i})^*$, we conclude that

$$\eta(B)_{i,k} = \delta_{i,k} \sum_{j=1}^d A_{i,j}^2 B_{j,j}.$$

Therefore $\eta(B) \in D_d(\mathbb{C})$ and the claimed expression holds. \square

The previous lemma, in notation of Theorem 3, shows that $T_{z\mathbf{I}}(D) \in D_d(\mathbb{C})$ for all $D \in D_d(\mathbb{C})$. This easily implies the following corollary.

Corollary 4. Let \mathbf{H} be a semicircular mixture, then $G_{\mathbf{H}}(z\mathbf{I}) \in D_d(\mathbb{C})$ and thus $G_{\mathbf{H}}(z\mathbf{I})_{i,i} \in \mathbb{C}^-$ for all $1 \leq i \leq d$.

It is important to notice that the following corollaries are weaker than the analysis done in [1], however they are enough to prove Theorem 1 so we include them for completeness.

Lemma 2. Let $z \in \mathbb{C}^+$, then $\|G_{\mathbf{H}}(z\mathbf{I})\|_1 \leq d\Im(z)^{-1}$.

Proof. Let $z \in \mathbb{C}^+$ be fixed and $i \in \{1, \dots, d\}$. By Lemma 1 and the fixed point equation (5) we have

$$\begin{aligned}1 &= (z - \eta(G_{\mathbf{H}}(z\mathbf{I}))_{i,i}) G_{\mathbf{H}}(z\mathbf{I})_{i,i} \\ &= \left(z - \sum_{j=1}^d A_{i,j}^2 G_{\mathbf{H}}(z\mathbf{I})_{j,j} \right) G_{\mathbf{H}}(z\mathbf{I})_{i,i}.\end{aligned}$$

By Corollary 4 we have that $\Im G_{\mathbf{H}}(z\mathbf{I})_{j,j} \leq 0$ for all $1 \leq j \leq d$, therefore

$$\begin{aligned}|G_{\mathbf{H}}(z\mathbf{I})_{i,i}| &= \left| z - \sum_{j=1}^d A_{i,j}^2 G_{\mathbf{H}}(z\mathbf{I})_{j,j} \right|^{-1} \\ &\leq \Im \left(z - \sum_{j=1}^d A_{i,j}^2 G_{\mathbf{H}}(z\mathbf{I})_{j,j} \right)^{-1} \\ &\leq \Im(z)^{-1},\end{aligned}$$

from which the result follows. \square

The proof of the following lemma follows the same lines as the previous one.

Lemma 3. Let $D \in M_d(\mathbb{C})^-$, then $\|T_{zI}(D)\|_1 \leq d\Im(z)^{-1}$.

Now we are ready to proof Theorem 1.

Proof of Theorem 1. By Theorem 3, for each $\omega \in \Omega$ we have that \mathbf{H} satisfies $G_{\mathbf{H}}(zI) = \lim_{n \rightarrow \infty} T_{zI}^{\circ n}(iI)$. By the definition of \mathcal{G} we have that $\mathcal{G}_{\mathbf{H}}(zI) = \mathbb{E}(G_{\mathbf{H}}(zI)) = \mathbb{E}(\lim_{n \rightarrow \infty} T_{zI}^{\circ n}(iI))$, so by the previous lemmas and the dominated convergence theorem,

$$\mathcal{G}_{\mathbf{H}}(zI) = \lim_{n \rightarrow \infty} \mathbb{E}(T_{zI}^{\circ n}(iI)),$$

as required. \square

Proof of Corollary 1. Let $\omega \in \Omega$ fixed. Observe that, for $z \in \mathbb{C}^+$, $\eta(zI) = KzI$. Thus the fixed point equation (5) implies that $KG_{\mathbf{H}}(zI)^2 - zG_{\mathbf{H}}(zI) + I = 0$. Equivalently, we have that $G_{\mathbf{H}}(zI) = \frac{z - \sqrt{z^2 - 4K}}{2K}I$ and therefore $\mathcal{G}_{\mathbf{H}}(zI) = \mathbb{E}(G_{\mathbf{H}}(zI)) = \mathbb{E}\left(\frac{z - \sqrt{z^2 - 4K}}{2K}\right)I$ as claimed. \square

Finally, we prove the central limit theorem, but first we have to prove the following lemma.

Lemma 4. Let $\mathbf{Z} \in L^{\infty-}(\Omega, \mathcal{F}, \mathbb{P}) \otimes M_d(\mathcal{A})$ be a selfadjoint centered random operator-valued matrix. If $\mathbf{Z}^{(1)}$ and $\mathbf{Z}^{(2)}$ are independent and free over $M_d(\mathbb{C})$ copies of \mathbf{Z} , then

$$\mathbb{E}\left(E\left(\mathbf{Z}^{(1)}\mathbf{Z}^{(2)}B\mathbf{Z}^{(2)}\mathbf{Z}^{(1)}\right)\right) = \hat{\eta}(\hat{\eta}(B))$$

where

$$\hat{\eta}(B) = \mathbb{E}(E(\mathbf{Z}B\mathbf{Z})).$$

Proof. Since $\mathbf{Z}^{(1)}$ and $\mathbf{Z}^{(2)}$ are free over $M_d(\mathbb{C})$,

$$\mathbb{E}\left(E\left(\mathbf{Z}^{(1)}\mathbf{Z}^{(2)}B\mathbf{Z}^{(2)}\mathbf{Z}^{(1)}\right)\right) = \mathbb{E}\left(E\left(\mathbf{Z}^{(1)}E\left(\mathbf{Z}^{(2)}B\mathbf{Z}^{(2)}\right)\mathbf{Z}^{(1)}\right)\right).$$

By the tower property of the conditional expectation

$$\begin{aligned} \mathbb{E}\left(E\left(\mathbf{Z}^{(1)}\mathbf{Z}^{(2)}B\mathbf{Z}^{(2)}\mathbf{Z}^{(1)}\right)\right) &= \mathbb{E}\left(\mathbb{E}\left(E\left(\mathbf{Z}^{(1)}E\left(\mathbf{Z}^{(2)}B\mathbf{Z}^{(2)}\right)\mathbf{Z}^{(1)}\right)\middle|\mathbf{Z}^{(1)}\right)\right) \\ &= \mathbb{E}\left(E\left(\mathbf{Z}^{(1)}\mathbb{E}\left(E\left(\mathbf{Z}^{(2)}B\mathbf{Z}^{(2)}\right)\middle|\mathbf{Z}^{(1)}\right)\mathbf{Z}^{(1)}\right)\right), \end{aligned}$$

where the last equation uses that conditional expectation and E commute (see Section 1). By independence,

$$\begin{aligned} \mathbb{E}\left(E\left(\mathbf{Z}^{(1)}\mathbf{Z}^{(2)}B\mathbf{Z}^{(2)}\mathbf{Z}^{(1)}\right)\right) &= \mathbb{E}\left(E\left(\mathbf{Z}^{(1)}\mathbb{E}\left(E\left(\mathbf{Z}^{(2)}B\mathbf{Z}^{(2)}\right)\right)\mathbf{Z}^{(1)}\right)\right) \\ &= \hat{\eta}(\hat{\eta}(B)), \end{aligned}$$

as required. \square

This random operator-valued version of the pair cancellation lemma allows us to prove the central limit theorem for random operator-valued matrices *mutatis mutandis* as in the operator-valued case.

Proof of Theorem 2. Let $m \in \mathbb{N}$ be fixed, then

$$\mathbb{E}(E(\mathbf{S}_N^m)) = N^{-m/2} \sum_{i_1, \dots, i_m=1}^N \mathbb{E}\left(E\left(\mathbf{Z}^{(i_1)} \dots \mathbf{Z}^{(i_m)}\right)\right).$$

As in the scalar and operator-valued cases, see [10] and [11] respectively, the independence, freeness and identically distributed assumptions imply that the value of $\mathbb{E}\left(E\left(\mathbf{Z}^{(i_1)} \dots \mathbf{Z}^{(i_m)}\right)\right)$ depends on i_1, \dots, i_m by means of which indices are equal and which are different. Let $\mathcal{P}(m)$ denote the set of all partitions of the integers $\{1, \dots, m\}$. For $\pi \in \mathcal{P}(m)$, we denote by B_π^N the number of tuples (i_1, \dots, i_m) of indices such that $1 \leq i_1, \dots, i_m \leq N$ and $i_j = i_k$ if and only if j and k belong to the same block in π . Also, we denote by C_π the value of $\mathbb{E}\left(E\left(\mathbf{Z}^{(i_1)} \dots \mathbf{Z}^{(i_m)}\right)\right)$ where (i_1, \dots, i_m) satisfies that $i_j = i_k$ if and only if j and k belong to the same block in π . We have then

$$\mathbb{E}(E(\mathbf{S}_N^m)) = N^{-m/2} \sum_{\pi \in \mathcal{P}(m)} B_\pi^N C_\pi.$$

As in the aforementioned references, we can analyze $\mathcal{P}(m)$ in four groups: partitions with singletons, non-crossing pairings, crossing pairings and *the rest*. Applying E , the freeness and centeredness assumptions imply that partitions with singletons and crossing pairings vanish. A simple combinatorial analysis shows that *the rest* does not contribute asymptotically. Therefore, just the non-crossing partitions contribute asymptotically, with $N^{-m/2} B_\pi^N \rightarrow 1$ as $N \rightarrow \infty$, and the previous pair cancellation lemma then leads to

$$\mathbb{E}(E(\mathbf{S}_N^m)) = \sum_{\pi \in \text{NC}_2(m)} \hat{\kappa}_\pi(\mathbf{Z}) + o(1).$$

This establishes the desired convergence. \square

It is important to notice that the independence were used only to apply the pair cancellation lemma (Lemma 4). Since the properties of E were used most of the time, it is natural then to expect that the limiting $M_d(\mathbb{C})$ -moments of the CLT limit have the same structure as in the operator-valued case.

Proof of Corollary 3. The hypothesis of the CLT are clearly satisfied and an easy computation shows that $\hat{\eta}_{A \circ \mathbf{X}}(B)_{i,j} = \sum_{i_1, i_2=1}^d B_{i_1, i_2} \mathbb{E}(A_{i, i_1} A_{i_2, j}) \varphi(X_{i, i_1} X_{i_2, j})$, which establishes equation (3). \square

References

- [1] HELTON, J. W., RASHIDI FAR, R. AND SPEICHER, R. (2007). Operator-valued semicircular elements: solving a quadratic matrix equation with positivity constraints. *International Mathematics Research Notices*, 15 pages.
- [2] BELINSCHI, S., MAI, T. AND SPEICHER, R. (2013). Analytic subordination theory of operator-valued free additive convolution and the solution of a general random matrix problem. *arXiv:1303.3196*.
- [3] BELINSCHI, S., SPEICHER, R., TREILHARD, J. AND VARGAS, C. (To appear). Operator-valued free multiplicative convolution: analytic subordination theory and applications to random matrix theory. *International Mathematics Research Notices*, 26 pages.
- [4] FAR, R., ORABY, T., BRYC W. AND SPEICHER R. (2008). On slow-fading MIMO systems with nonseparable correlation. *IEEE Trans. on Information Theory* 54, 544–553.
- [5] SPEICHER, R., VARGAS, C. AND MAI, T. (2012). Free deterministic equivalents, rectangular random matrix models and operator-valued free probability theory. *Random Matrices: Theory and Applications* 1.
- [6] DIAZ, M. AND PÉREZ-ABREU, V. (2014). Random matrix systems with block-based behavior and operator-valued models. *arXiv:1404.4420*.
- [7] LARSSON, E., EDFORS, O. AND MARZETTA T. (2014) Massive MIMO for next generation wireless systems. *IEEE Communications Magazine* Feb, 186–195.
- [8] HIAI, F. AND PETZ, D. (2000). *The Semicircle Law, Free Random Variables and Entropy*. American Mathematical Society, United States.
- [9] MINGO, J. AND SPEICHER, R. (To appear). *Free Probability and Random Matrices*.
- [10] NICA, A. AND SPEICHER, R. (2006). *Lectures on the Combinatorics of Free Probability*. Cambridge University Press, United Kingdom.
- [11] SPEICHER, R. (1998). *Combinatorial Theory of the Free Product with Amalgamation and Operator-Valued Free Probability Theory*. American Mathematical Society, United States of America.